Linear Algebra

Lecture 5 - Operations on Linear Transformations

Oskar Kędzierski

4 November 2024

Sum and Scalar Multiplication

Proposition

Let V, W be vector spaces. Let $\varphi, \psi : V \longrightarrow W$ be linear transformations and let $\alpha \in \mathbb{R}$. The transformation $\varphi + \psi : V \longrightarrow W$, defined by $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$ for $v \in V$, and the transformation $\alpha\varphi$ defined by $(\alpha\varphi)(v) = \alpha\varphi(v)$ are linear. The transformation $\varphi + \psi$ is called the **sum** *of φ and ψ and $\alpha\varphi$ is called the* **product** *of the transformation φ with scalar α.*

Sum and Scalar Multiplication

Proposition

Let V, W be vector spaces. Let $\varphi, \psi : V \longrightarrow W$ be linear transformations and let $\alpha \in \mathbb{R}$. The transformation $\varphi + \psi : V \longrightarrow W$, defined by $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$ for $v \in V$, and the transformation $\alpha\varphi$ defined by $(\alpha\varphi)(v) = \alpha\varphi(v)$ are linear. The transformation $\varphi + \psi$ is called the **sum** *of φ and ψ and $\alpha\varphi$ is called the* **product** *of the transformation φ with scalar α.*

Example

Let $\varphi, \psi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be given by
$\varphi((x_1, x_2, x_3)) = (x_1 + 2x_2 - x_3, x_1 + 2x_2 + x_3)$ and
$\psi((x_1, x_2, x_3)) = (-x_1 + x_2 + x_3, 3x_1 - 2x_2 + x_3)$. Then
$(\varphi + \psi)((x_1, x_2, x_3)) = (3x_2, 4x_1 + 2x_3)$ and
$(2\varphi)((x_1, x_2, x_3)) = (2x_1 + 4x_2 - 2x_3, 2x_1 + 4x_2 + 2x_3)$ (for $\alpha = 2$).

Composition

Proposition

Let U, V, W be vectors spaces and let $\varphi : U \longrightarrow V$, $\psi : V \longrightarrow W$ be linear transformations. The transformation

$$\psi \circ \varphi \colon U \longrightarrow W,$$

given by

$$(\psi \circ \varphi)(v) = \psi(\varphi(v)))$$

for $v \in U$, is linear. It is called the **composition** *of ψ with φ.*

Example

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ and $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be linear transformations given by $\varphi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, -x_1 + 3x_2 - x_3)$ and $\psi((y_1, y_2)) = (y_1 - y_2, y_1 + 2y_2))$. Then $(\psi \circ \varphi)((x_1, x_2, x_3)) = \psi((x_1 - x_2 + 2x_3, -x_1 + 3x_2 - x_3)) =$ $((x_1 - x_2 + 2x_3) - (-x_1 + 3x_2 - x_3), (x_1 - x_2 + 2x_3) + 2(-x_1 + 3x_2 - x_3)) =$ $(2x_1 - 4x_2 + 3x_3, -x_1 + 5x_2)$.

Operations on Matrices

Definition

Let $A, B \in M(m \times n; \mathbb{R}), \alpha \in \mathbb{R},\ A = [a_{ij}], B = [b_{ij}]$. The **sum** of matrices A and B is the matrix $A + B = [a_{ij} + b_{ij}]$. The **product** of matrix A by scalar α is the matrix $\alpha A = [\alpha a_{ij}]$.

Operations on Matrices

Definition

Let $A, B \in M(m \times n; \mathbb{R}), \alpha \in \mathbb{R}, \; A = [a_{ij}], B = [b_{ij}]$. The **sum** of matrices A and B is the matrix $A + B = [a_{ij} + b_{ij}]$. The **product** of matrix A by scalar α is the matrix $\alpha A = [\alpha a_{ij}]$.

Example

Let $\alpha = 2$ and let $A, B \in M(2 \times 3; \mathbb{R})$ be given by

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} -1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix} 0 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \; \alpha A = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 2 & 0 \end{bmatrix}.$$

Matrix Multiplication

Definition

Let $A \in M(m \times n; \mathbb{R})$ and let $B \in M(n \times l; \mathbb{R})$. The **matrix product** of A by B is the matrix $AB = [c_{ij}] \in M(m \times l; \mathbb{R})$ where $c_{ij} = \sum_{s=1}^{n} a_{is} b_{sj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, l$.

Matrix Multiplication

Definition

Let $A \in M(m \times n; \mathbb{R})$ and let $B \in M(n \times l; \mathbb{R})$. The **matrix product** of A by B is the matrix $AB = [c_{ij}] \in M(m \times l; \mathbb{R})$ where $c_{ij} = \sum_{s=1}^{n} a_{is} b_{sj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, l$.

In particular, if $R_i = \begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix} \in M(1 \times n; \mathbb{R})$ is the $i-$th row of matrix A and $C_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \in M(n \times 1; \mathbb{R})$ is the $j-$th column of matrix B then $R_i C_j = \begin{bmatrix} a_{i1} b_{1j} + \ldots + a_{in} b_{nj} \end{bmatrix}$ is a 1×1 matrix which can be identified with a real number.

Matrix Multiplication (continued)

Using this identification we can write

$$AB = \begin{bmatrix} R_1 C_1 & R_1 C_2 & \dots & R_1 C_l \\ R_2 C_1 & R_2 C_2 & \dots & R_2 C_l \\ \vdots & \vdots & \ddots & \vdots \\ R_m C_1 & R_m C_2 & \dots & R_m C_l \end{bmatrix}.$$

Matrix Multiplication (continued)

Using this identification we can write

$$AB = \begin{bmatrix} R_1 C_1 & R_1 C_2 & \dots & R_1 C_l \\ R_2 C_1 & R_2 C_2 & \dots & R_2 C_l \\ \vdots & \vdots & \ddots & \vdots \\ R_m C_1 & R_m C_2 & \dots & R_m C_l \end{bmatrix}.$$

For example

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} = \text{the first column of } A$$

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix} =$$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \end{bmatrix} = \text{the first row of } B$$

and so on.

Example

Let $A \in M(3 \times 2; \mathbb{R})$ and $B \in M(2 \times 2; \mathbb{R})$ be given by

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} \begin{bmatrix} C_1 & C_2 \end{bmatrix} = \begin{bmatrix} R_1 C_1 & R_1 C_2 \\ R_2 C_1 & R_2 C_2 \\ R_3 C_1 & R_3 C_3 \end{bmatrix} = \begin{bmatrix} 3 & -3 \\ 5 & -4 \\ 0 & -3 \end{bmatrix}.$$

The first column of AB is the sum of columns of A and the second one is the first column of A minus twice the second column of A.

Warning

The matrix multiplication is, in general, not commutative. For example

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

but

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Operations on Linear Transformations and Matrices

Theorem (Addition)

Let V, W be vector spaces and let $\varphi, \psi : V \longrightarrow W$ be linear transformations. Let $\mathcal{A}, \mathcal{B}$ be bases of V and W respectively. Then $M(\varphi + \psi)_{\mathcal{A}}^{\mathcal{B}} = M(\varphi)_{\mathcal{A}}^{\mathcal{B}} + M(\psi)_{\mathcal{A}}^{\mathcal{B}}$.

Theorem (Composition and multiplication)

Let U, V, W be vectors spaces and let $\varphi : U \longrightarrow V$, $\psi : V \longrightarrow W$ be linear transformations. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be the bases of U, V and W, respectively. Then $M(\psi \circ \varphi)_{\mathcal{A}}^{\mathcal{C}} = M(\psi)_{\mathcal{B}}^{\mathcal{C}} M(\varphi)_{\mathcal{A}}^{\mathcal{B}}$.

Example (continued)

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ and $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be linear transformations given by $\varphi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, -x_1 + 3x_2 - x_3)$ and $\psi((y_1, y_2)) = (y_1 - y_2, y_1 + 2y_2)$. Recall that $(\psi \circ \varphi)((x_1, x_2, x_3)) = (2x_1 - 4x_2 + 3x_3, -x_1 + 5x_2)$.

Example (continued)

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ and $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be linear transformations given by $\varphi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, -x_1 + 3x_2 - x_3)$ and $\psi((y_1, y_2)) = (y_1 - y_2, y_1 + 2y_2)$. Recall that $(\psi \circ \varphi)((x_1, x_2, x_3)) = (2x_1 - 4x_2 + 3x_3, -x_1 + 5x_2)$. We will compute this again, using matrix multiplication.

Example (continued)

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ and $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be linear transformations given by $\varphi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, -x_1 + 3x_2 - x_3)$ and $\psi((y_1, y_2)) = (y_1 - y_2, y_1 + 2y_2)$. Recall that $(\psi \circ \varphi)((x_1, x_2, x_3)) = (2x_1 - 4x_2 + 3x_3, -x_1 + 5x_2)$. We will compute this again, using matrix multiplication. Let $\mathcal{A}$ be the standard basis in $\mathbb{R}^3$ and let $\mathcal{B} = \mathcal{C}$ be the standard basis in $\mathbb{R}^2$. Then

$$M(\psi \circ \varphi)_{\mathcal{A}}^{\mathcal{C}} = M(\psi)_{\mathcal{B}}^{\mathcal{C}} M(\varphi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 3 & -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 & -4 & 3 \\ -1 & 5 & 0 \end{bmatrix}.$$

This agrees with the formula of $\psi \circ \varphi$.

Applications

Proposition

Let V, W be vector spaces and let $\varphi : V \longrightarrow W$ be a linear transformation. Let $\mathcal{A} = (v_1, \ldots, v_n)$ be an ordered basis of V and let $\mathcal{B} = (w_1, \ldots, w_m)$ be an ordered basis of W. For any vector $v \in V$ let $\alpha_1, \ldots, \alpha_n$ be the coordinates of v relative to the basis $\mathcal{A}$ and let $\beta_1, \ldots, \beta_m$ be the coordinates of $\varphi(v)$ relative to the basis $\mathcal{B}$, that is $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ and $\varphi(v) = \beta_1 w_1 + \ldots + \beta_m w_m$. Then

$$M(\varphi)_{\mathcal{A}}^{\mathcal{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}.$$

Example

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformations given by $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis in $\mathbb{R}^2$ and let $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$.

Example

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformations given by $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis in $\mathbb{R}^2$ and let
$\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$.
We check immediately that

$$\psi(1,2) = (-1,5) = 4(1,0) - 5(1,-1),$$

$$\psi(0,1) = (-1,2) = 1(1,0) - 2(1,-1).$$

Therefore

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \quad M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}.$$

Example

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformations given by $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis in $\mathbb{R}^2$ and let $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$. We check immediately that

$$\psi(1,2) = (-1,5) = 4(1,0) - 5(1,-1),$$

$$\psi(0,1) = (-1,2) = 1(1,0) - 2(1,-1).$$

Therefore

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}.$$

Pick, say, $v = (1,1)$. Since $v = 1(1,2) - 1(0,1)$, the coordinates of v relative to $\mathcal{A}$ are $1, -1$. Since $\psi(v) = (0,3) = 3(1,0) - 3(1,-1)$, the coordinates of $\psi(v)$ relative to $\mathcal{B}$ are $3, -3$.

Example (continued)

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \quad M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}.$$

Example (continued)

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}.$$

the coordinates of $v = (1, 1)$ relative to the basis $\mathcal{A}$ are $1, -1$

the coordinates of $\psi(v) = (0, 3)$ relative to the basis $\mathcal{B}$ are $3, -3$

Example (continued)

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}.$$

the coordinates of $v = (1,1)$ relative to the basis $\mathcal{A}$ are $1, -1$

the coordinates of $\psi(v) = (0,3)$ relative to the basis $\mathcal{B}$ are $3, -3$

$$M(\psi)_{st}^{st} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$M(\psi)_{\mathcal{A}}^{\mathcal{B}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}.$$

Applications (continued)

Let V be a vector space. The function $\mathrm{id}_V : V \longrightarrow V$ given by $id_V(v) = v$ for any $v \in V$ is a linear transformation called **the identity**.

Applications (continued)

Let V be a vector space. The function $\mathrm{id}_V : V \longrightarrow V$ given by $id_V(v) = v$ for any $v \in V$ is a linear transformation called **the identity**.

Corollary

Let $\mathcal{A} = (v_1, \ldots, v_n)$ and $\mathcal{B} = (w_1, \ldots, w_n)$ be two ordered bases of V. For any $v \in V$ let $\alpha_1, \ldots, \alpha_n$ be the coordinates of v relative to the basis $\mathcal{A}$ and let $\beta_1, \ldots, \beta_n$ be the coordinates of v relative to the basis $\mathcal{B}$. Then

$$M(\mathrm{id}_V)_{\mathcal{A}}^{\mathcal{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Applications (continued)

Let V be a vector space. The function $\mathrm{id}_V : V \longrightarrow V$ given by $id_V(v) = v$ for any $v \in V$ is a linear transformation called **the identity**.

Corollary

Let $\mathcal{A} = (v_1, \ldots, v_n)$ and $\mathcal{B} = (w_1, \ldots, w_n)$ be two ordered bases of V. For any $v \in V$ let $\alpha_1, \ldots, \alpha_n$ be the coordinates of v relative to the basis $\mathcal{A}$ and let $\beta_1, \ldots, \beta_n$ be the coordinates of v relative to the basis $\mathcal{B}$. Then

$$M(\mathrm{id}_V)_{\mathcal{A}}^{\mathcal{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

The matrix $M(\mathrm{id}_V)_{\mathcal{A}}^{\mathcal{B}}$ is called a **change-of-coordinates matrix**.

Applications (continued)

Proposition

Let V, W be vector spaces and let $\varphi : V \longrightarrow W$ be a linear transformation. Let $\mathcal{A}, \mathcal{A}'$ be (ordered) bases of V and let $\mathcal{B}, \mathcal{B}'$ be (ordered) bases of W. Then

$$M(\varphi)^{\mathcal{B}'}_{\mathcal{A}'} = M(\mathrm{id}_W)^{\mathcal{B}'}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}} M(\mathrm{id}_V)^{\mathcal{A}}_{\mathcal{A}'}.$$

Proof.

This follows directly from the fact that $\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V = \varphi$ and the formula relating composition of linear transformations with matrix multiplication. $\square$

Example (continued)

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation given by the formula $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1, 0), (0, 1))$ be the standard basis of $\mathbb{R}^2$ and let $\mathcal{A} = ((1, 2), (0, 1))$, $\mathcal{B} = ((1, 0), (1, -1))$ be other two bases of $\mathbb{R}^2$.

Example (continued)

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation given by the formula $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis of $\mathbb{R}^2$ and let $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$.We have already checked that

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}$$

Example (continued)

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation given by the formula $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis of $\mathbb{R}^2$ and let $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$.We have already checked that

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}$$

Let check this again using the previous Proposition. It says that

$$M(\psi)_{\mathcal{A}}^{\mathcal{B}} = M(\mathrm{id}_{\mathbb{R}^2})_{st}^{\mathcal{B}} M(\psi)_{st}^{st} M(\mathrm{id}_{\mathbb{R}^2})_{\mathcal{A}}^{st}$$

Example (continued)

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation given by the formula $\psi((x_1, x_2)) = (x_1 - x_2, x_1 + 2x_2)$. Let $st = ((1,0),(0,1))$ be the standard basis of $\mathbb{R}^2$ and let $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$ be other two bases of $\mathbb{R}^2$.We have already checked that

$$M(\psi)_{st}^{st} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \; M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix}$$

Let check this again using the previous Proposition. It says that

$$M(\psi)_{\mathcal{A}}^{\mathcal{B}} = M(\mathrm{id}_{\mathbb{R}^2})_{st}^{\mathcal{B}} M(\psi)_{st}^{st} M(\mathrm{id}_{\mathbb{R}^2})_{\mathcal{A}}^{st}$$

We need to compute $M(\mathrm{id}_{\mathbb{R}^2})_{st}^{\mathcal{B}}$ and $M(\mathrm{id}_{\mathbb{R}^2})_{\mathcal{A}}^{st}$.

Example (continued)

We need to compute $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st}$ and $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}}$. Recall that $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$.

Example (continued)

We need to compute $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st}$ and $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}}$. Recall that $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$. Since

$$\mathrm{id}((1,2)) = 1(1,0) + 2(0,1),$$

$$\mathrm{id}(0,1) = 0(1,0) + 1(0,1),$$

we have $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$.

Example (continued)

We need to compute $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st}$ and $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}}$. Recall that $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$. Since

$$\mathrm{id}((1,2)) = 1(1,0) + 2(0,1),$$

$$\mathrm{id}(0,1) = 0(1,0) + 1(0,1),$$

we have $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$. Since

$$\mathrm{id}((1,0)) = 1(1,0) + 0(1,-1),$$

$$\mathrm{id}((0,1)) = 1(1,0) - 1(1,-1),$$

we have $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$.

Example (continued)

We need to compute $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st}$ and $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}}$. Recall that $\mathcal{A} = ((1,2),(0,1))$, $\mathcal{B} = ((1,0),(1,-1))$. Since

$$\mathrm{id}((1,2)) = 1(1,0) + 2(0,1),$$

$$\mathrm{id}(0,1) = 0(1,0) + 1(0,1),$$

we have $M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$. Since

$$\mathrm{id}((1,0)) = 1(1,0) + 0(1,-1),$$

$$\mathrm{id}((0,1)) = 1(1,0) - 1(1,-1),$$

we have $M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$. Using $M(\psi)^{\mathcal{B}}_{\mathcal{A}} = M(\mathrm{id}_{\mathbb{R}^2})^{\mathcal{B}}_{st} M(\psi)^{st}_{st} M(\mathrm{id}_{\mathbb{R}^2})^{st}_{\mathcal{A}}$ one can check that

$$\begin{bmatrix} 4 & 1 \\ -5 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

Elementary Matrices

Fix $\alpha \in \mathbb{R}$, $n > 0$ and define the following matrices $D_{i,\alpha} = [d_{kl}], L_{ij} = [\ell_{kl}], T_{ij} = [t_{kl}] \in M(n \times n; \mathbb{R})$ as follows

Elementary Matrices

Fix $\alpha \in \mathbb{R}, \;\; n > 0$ and define the following matrices $D_{i,\alpha} = [d_{kl}], L_{ij} = [\ell_{kl}], T_{ij} = [t_{kl}] \in M(n \times n; \mathbb{R})$ as follows

i) $d_{kk} = 1$ for $k \neq i, \;\; d_{ii} = \alpha, \;\; d_{kl} = 0$ elsewhere,

Elementary Matrices

Fix $\alpha \in \mathbb{R}$, $n > 0$ and define the following matrices $D_{i,\alpha} = [d_{kl}], L_{ij} = [\ell_{kl}], T_{ij} = [t_{kl}] \in M(n \times n; \mathbb{R})$ as follows

i) $d_{kk} = 1$ for $k \neq i$, $d_{ii} = \alpha$, $d_{kl} = 0$ elsewhere,

ii) $\ell_{kk} = 1$ for $k = 1, \ldots, n$, $\ell_{ij} = 1$, $\ell_{kl} = 0$ elsewhere,

Elementary Matrices

Fix $\alpha \in \mathbb{R}$, $n > 0$ and define the following matrices $D_{i,\alpha} = [d_{kl}], L_{ij} = [\ell_{kl}], T_{ij} = [t_{kl}] \in M(n \times n; \mathbb{R})$ as follows

i) $d_{kk} = 1$ for $k \neq i$, $d_{ii} = \alpha$, $d_{kl} = 0$ elsewhere,

ii) $\ell_{kk} = 1$ for $k = 1, \ldots, n$, $\ell_{ij} = 1$, $\ell_{kl} = 0$ elsewhere,

iii) $t_{kk} = 1$ for $k \notin \{i, j\}$, $t_{ij} = t_{ji} = 1$, $t_{kl} = 0$ elsewhere.

Elementary Matrices (continued)

$$D_{i,\alpha} = \begin{array}{cc} & \begin{matrix} & i & & & & & \end{matrix} \\ \begin{matrix} \\ i \\ \\ \\ \\ \\ \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \alpha & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array}, \qquad L_{ij} = \begin{array}{cc} & \begin{matrix} & & & j & & & \end{matrix} \\ \begin{matrix} \\ i \\ \\ \\ \\ \\ \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \end{array},$$

$$T_{ij} = \begin{array}{cc} & \begin{matrix} & & i & & & & j & \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ \\ \\ j \\ \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$*. Then*

i) $D_{i,\alpha}A =$ *matrix* A *with the* i*-th row multiplied by* α,

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. *Then*

i) $D_{i,\alpha}A =$ *matrix* A *with the* i*-th row multiplied by* α,

ii) $L_{ij}A =$ *matrix* A *with the* j*-th row added to the* i*-th row*,

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. *Then*

i) $D_{i,\alpha}A =$ *matrix* A *with the* i*-th row multiplied by* α,

ii) $L_{ij}A =$ *matrix* A *with the* j*-th row added to the* i*-th row*,

iii) $T_{ij}A =$ *matrix* A *with the* i*-th and* j*-th rows swapped*,

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $D_{i,\alpha} A =$ *matrix A with the i-th row multiplied by α,*

ii) $L_{ij} A =$ *matrix A with the j-th row added to the i-th row,*

iii) $T_{ij} A =$ *matrix A with the i-th and j-th rows swapped,*

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $D_{i,\alpha}A =$ *matrix A with the i-th row multiplied by α,*

ii) $L_{ij}A =$ *matrix A with the j-th row added to the i-th row,*

iii) $T_{ij}A =$ *matrix A with the i-th and j-th rows swapped,*

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $D_{i,\alpha}A =$ *matrix A with the i-th row multiplied by α,*

ii) $L_{ij}A =$ *matrix A with the j-th row added to the i-th row,*

iii) $T_{ij}A =$ *matrix A with the i-th and j-th rows swapped,*

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $AD_{i,\alpha} =$ *matrix A with the i-th column multiplied by α,*

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. *Then*

i) $D_{i,\alpha}A =$ *matrix* A *with the* i*-th row multiplied by* α,

ii) $L_{ij}A =$ *matrix* A *with the* j*-th row added to the* i*-th row*,

iii) $T_{ij}A =$ *matrix* A *with the* i*-th and* j*-th rows swapped*,

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Proposition

Let $A \in M(n \times m; \mathbb{R})$. *Then*

i) $AD_{i,\alpha} =$ *matrix* A *with the* i*-th column multiplied by* α,

ii) $AL_{ij} =$ *matrix* A *with the* i*-th column added to the* j*-th one*,

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $D_{i,\alpha}A =$ *matrix A with the i-th row multiplied by α,*

ii) $L_{ij}A =$ *matrix A with the j-th row added to the i-th row,*

iii) $T_{ij}A =$ *matrix A with the i-th and j-th rows swapped,*

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $AD_{i,\alpha} =$ *matrix A with the i-th column multiplied by α,*

ii) $AL_{ij} =$ *matrix A with the i-th column added to the j-th one,*

iii) $AT_{ij} =$ *matrix A with the i-th and j-th columns swapped,*

Elementary Matrices (continued)

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $D_{i,\alpha}A =$ *matrix A with the i-th row multiplied by α,*

ii) $L_{ij}A =$ *matrix A with the j-th row added to the i-th row,*

iii) $T_{ij}A =$ *matrix A with the i-th and j-th rows swapped,*

that is, elementary row operations correspond to multiplication by elementary matrices from the left.

Proposition

Let $A \in M(n \times m; \mathbb{R})$. Then

i) $AD_{i,\alpha} =$ *matrix A with the i-th column multiplied by α,*

ii) $AL_{ij} =$ *matrix A with the i-th column added to the j-th one,*

iii) $AT_{ij} =$ *matrix A with the i-th and j-th columns swapped,*

that is, elementary column operations correspond to multiplication by elementary matrices from the right.

Matrix Multiplication is Associative

Proposition

For any matrices
$A \in M(m \times n; \mathbb{R}), B \in M(n \times l; \mathbb{R}), C \in M(l \times k; \mathbb{R})$

$$(AB)C = A(BC).$$

Matrix Multiplication is Associative

Proposition

For any matrices
$A \in M(m \times n; \mathbb{R}), B \in M(n \times l; \mathbb{R}), C \in M(l \times k; \mathbb{R})$

$$(AB)C = A(BC).$$

Proof.

Let $AB = [f_{ij}] \in M(m \times l; \mathbb{R}), BC = [g_{ij}] \in M(n \times k; \mathbb{R})$. Then

Matrix Multiplication is Associative

Proposition

For any matrices
$A \in M(m \times n; \mathbb{R}), B \in M(n \times l; \mathbb{R}), C \in M(l \times k; \mathbb{R})$

$$(AB)C = A(BC).$$

Proof.

Let $AB = [f_{ij}] \in M(m \times l; \mathbb{R}), BC = [g_{ij}] \in M(n \times k; \mathbb{R})$. Then

$$f_{ir} = \sum_{s=1}^{n} a_{is} b_{sr},$$

$$g_{sj} = \sum_{r=1}^{l} b_{sr} c_{rj}.$$

Matrix Multiplication is Associative (continued)

Proof.

The entry in the i-th row and the j-th column of the matrix $(AB)C$ is equal to

$$\sum_{r=1}^{l} f_{ir} c_{rj} = \sum_{r=1}^{l} \left(\sum_{s=1}^{n} a_{is} b_{sr} \right) c_{rj} = \sum_{r=1}^{l} \sum_{s=1}^{n} a_{is} b_{sr} c_{rj}.$$

The entry in the i-th row and the j-th column of the matrix $A(BC)$ is equal to

$$\sum_{s=1}^{n} a_{is} g_{sj} = \sum_{s=1}^{n} a_{is} \left(\sum_{r=1}^{l} b_{sr} c_{rj} \right) = \sum_{s=1}^{n} \sum_{r=1}^{l} a_{is} b_{sr} c_{rj} =$$

$$= \sum_{r=1}^{l} \sum_{s=1}^{n} a_{is} b_{sr} c_{rj}.$$

$\square$

Coordinate Vector

Definition

Let V be a vector space and let $\mathcal{A} = (v_1, \ldots, v_n)$ be its ordered basis. For any $v \in V$ by $\left[v\right]_{\mathcal{A}}$ we denote the **coordinate vector** of v relative to $\mathcal{A}$, i.e. a n–by–1 matrix with coordinates of v relative to $\mathcal{A}$. In particular, if $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ then

$$\left[v\right]_{\mathcal{A}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Coordinate Vector

Definition

Let V be a vector space and let $\mathcal{A} = (v_1, \dots, v_n)$ be its ordered basis. For any $v \in V$ by $\left[v\right]_{\mathcal{A}}$ we denote the **coordinate vector** of v relative to $\mathcal{A}$, i.e. a n–by–1 matrix with coordinates of v relative to $\mathcal{A}$. In particular, if $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ then

$$\left[v\right]_{\mathcal{A}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Example

If $\mathcal{A} = ((1,1),(1,2)),\ v = (1,3)$ then

$$\left[v\right]_{st} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \text{ and } \left[v\right]_{\mathcal{A}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Image by the Matrix of Linear Transformation

Proposition

Let V, W be vector spaces and let $\varphi : V \longrightarrow W$ be a linear transformation. Let $\mathcal{A} = (v_1, \ldots, v_n)$ be an ordered basis of V and let $\mathcal{B} = (w_1, \ldots, w_m)$ be an ordered basis of W. For any vector $v \in V$ let $\alpha_1, \ldots, \alpha_n$ be the coordinates of v relative to the basis $\mathcal{A}$ and let $\beta_1, \ldots, \beta_m$ be the coordinates of $\varphi(v)$ relative to the basis $\mathcal{B}$, that is $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ and $\varphi(v) = \beta_1 w_1 + \ldots + \beta_m w_m$. Then

$$M(\varphi)^{\mathcal{B}}_{\mathcal{A}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$$

or equivalently

$$M(\varphi)^{\mathcal{B}}_{\mathcal{A}} [v]_{\mathcal{A}} = [\varphi(v)]_{\mathcal{B}}.$$

Image by the Matrix of Linear Transformation – Proof

$$M(\varphi)_{\mathcal{A}}^{\mathcal{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} =$$

$$= \alpha_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Image by the Matrix of Linear Transformation – Proof

$$M(\varphi)_{\mathcal{A}}^{\mathcal{B}}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}=$$

$$=\alpha_1\begin{bmatrix}a_{11}\\ a_{21}\\ \vdots\\ a_{m1}\end{bmatrix}+\alpha_2\begin{bmatrix}a_{12}\\ a_{22}\\ \vdots\\ a_{m2}\end{bmatrix}+\ldots+\alpha_n\begin{bmatrix}a_{1n}\\ a_{2n}\\ \vdots\\ a_{mn}\end{bmatrix}.$$

On the other hand

$$\alpha_1\varphi(v_1)=\alpha_1(a_{11}w_1+a_{21}w_2+\ldots+a_{m1}w_m),$$
$$\alpha_2\varphi(v_2)=\alpha_2(a_{12}w_1+a_{22}w_2+\ldots+a_{m2}w_m),$$
$$\vdots$$
$$\alpha_n\varphi(v_n)=\alpha_n(a_{1n}w_1+a_{2n}w_2+\ldots+a_{mn}w_m),$$

and

$$\varphi(v)=\varphi(\alpha_1v_1+\alpha_2v_2+\ldots+\alpha_nv_n)=\alpha_1\varphi(v_1)+\alpha_2\varphi(v_2)+\ldots+\alpha_n\varphi(v_n).$$

Matrix of the Sum of Linear Transformations

Proposition

Let V, W be vector spaces and let $\varphi, \psi : V \longrightarrow W$ be linear transformations. Let $\mathcal{A}, \mathcal{B}$ be bases of V and W respectively and $\alpha \in \mathbb{R}$ any scalar. Then $M(\varphi + \psi)^{\mathcal{B}}_{\mathcal{A}} = M(\varphi)^{\mathcal{B}}_{\mathcal{A}} + M(\psi)^{\mathcal{B}}_{\mathcal{A}}$ and $M(\alpha\varphi)^{\mathcal{B}}_{\mathcal{A}} = \alpha M(\varphi)^{\mathcal{B}}_{\mathcal{A}}$.

Proof.

Let $\mathcal{A} = (v_1, \dots, v_n)$ and $\mathcal{B} = (w_1, \dots, w_m)$ be the ordered bases of V and W, respectively. If

$$\begin{aligned} \varphi(v_i) &= a_{1i}w_1 + a_{2i}w_2 + \dots + a_{mi}w_m, \\ \psi(v_i) &= b_{1i}w_1 + b_{2i}w_2 + \dots + b_{mi}w_m, \end{aligned}$$

then

$$(\varphi + \psi)(v_i) = \varphi(v_i) + \psi(v_i) = (a_{1i} + b_{1i})w_1 + (a_{2i} + b_{2i})w_2 + \dots$$
$$\dots + (a_{mi} + b_{mi})w_m,$$
$$(\alpha\varphi)(v_i) = \alpha\varphi(v_i) = (\alpha a_{1i})w_1 + (\alpha a_{2i})w_2 + \dots + (\alpha a_{mi})w_m.$$

Matrix of the Composition of Linear Transformations

Proposition

Let U, V, W be vectors spaces and let $\varphi : U \longrightarrow V$, $\psi : V \longrightarrow W$ be linear transformations. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be the bases of U, V and W, respectively. Then $M(\psi \circ \varphi)^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}$.

Matrix of the Composition of Linear Transformations

Proposition

Let U, V, W be vectors spaces and let $\varphi : U \longrightarrow V$, $\psi : V \longrightarrow W$ be linear transformations. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be the bases of U, V and W, respectively. Then $M(\psi \circ \varphi)^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}$.

Proof.

By the Proposition on the image by matrix of a linear transformation

$$M(\varphi)^{\mathcal{B}}_{\mathcal{A}}\big[v\big]_{\mathcal{A}} = \big[\varphi(v)\big]_{\mathcal{B}},$$
$$M(\psi)^{\mathcal{C}}_{\mathcal{B}}\big[\varphi(v)\big]_{\mathcal{B}} = \big[\psi\left(\varphi(v)\right)\big]_{\mathcal{C}} = \big[(\psi \circ \varphi)(v)\big]_{\mathcal{C}},$$

that is, by associativity of matrix product

$$\big[(\psi \circ \varphi)(v)\big]_{\mathcal{C}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}}\left(M(\varphi)^{\mathcal{B}}_{\mathcal{A}}[v]_{\mathcal{A}}\right) = \left(M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}\right)\big[v\big]_{\mathcal{A}}.$$

Matrix of the Composition of Linear Transformations (continued)

Proof.

Therefore, for any $v \in V$

$$M(\psi \circ \varphi)^{\mathcal{C}}_{\mathcal{A}}\big[v\big]_{\mathcal{A}} = \big[(\varphi \circ \psi)(v)\big]_{\mathcal{C}} = \big(M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}\big)\,\big[v\big]_{\mathcal{A}}.$$

Substituting $v = v_i$ for $i = 1, \dots, n$ we see that matrices $M(\psi \circ \varphi)^{\mathcal{C}}_{\mathcal{A}}, M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}$ have the same columns, in particular

$$M(\psi \circ \varphi)^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} M(\varphi)^{\mathcal{B}}_{\mathcal{A}}.$$

□

Block Matrix

Definition

Matrix $A \in M(m \times n; \mathbb{R})$ is a **block matrix** if

$$A = \left[\begin{array}{c|c|c|c} \multicolumn{1}{c}{A_{11}} & \multicolumn{1}{c}{A_{12}} & \multicolumn{1}{c}{\cdots} & \multicolumn{1}{c}{A_{1q}} \\ \hline A_{21} & A_{22} & & A_{2q} \\ \hline \vdots & & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right],$$

where $A_{i,j} \in M(m_i \times n_j; \mathbb{R})$ and

$$m = m_1 + m_2 + \ldots + m_p,$$

$$n = n_1 + n_2 + \ldots + n_q.$$

Example

$$A = \left[\begin{array}{c|ccc} 1 & 2 & 2 & 6 \\ 1 & 2 & 2 & 5 \\ 1 & 1 & 2 & 8 \\ \hline 2 & 5 & 6 & 2 \end{array}\right].$$

Multiplication of Block Matrices

Proposition

If $A \in M(m \times n; \mathbb{R})$, $B \in M(n \times l; \mathbb{R})$ are block matrices, where

$$m = m_1 + \ldots + m_p, \quad n = n_1 + \ldots + n_q, \quad l = l_1 + \ldots + l_r,$$

$$A = [A_{ij}], \quad B = [B_{jk}],$$

for $i = 1, \ldots, p$, $j = 1, \ldots, q$, $k = 1, \ldots, r$, then $C = AB$ is a block matrix such that $C = [C_{ik}]$, where

$$C_{ik} = \sum_{j=1}^{q} A_{ij} B_{jk},$$

that is

$$C_{ik} = A_{i1}B_{1k} + A_{i2}B_{2k} + \ldots + A_{iq}B_{qk}.$$

Multiplication of Block Matrices

Proposition

If $A \in M(m \times n; \mathbb{R})$, $B \in M(n \times l; \mathbb{R})$ are block matrices, where

$$m = m_1 + \ldots + m_p, \quad n = n_1 + \ldots + n_q, \quad l = l_1 + \ldots + l_r,$$

$$A = [A_{ij}], \quad B = [B_{jk}],$$

for $i = 1, \ldots, p$, $j = 1, \ldots, q$, $k = 1, \ldots, r$, then $C = AB$ is a block matrix such that $C = [C_{ik}]$, where

$$C_{ik} = \sum_{j=1}^{q} A_{ij} B_{jk},$$

that is

$$C_{ik} = A_{i1} B_{1k} + A_{i2} B_{2k} + \ldots + A_{iq} B_{qk}.$$

Proof.

Exercise.

Multiplication of Block Matrices (continued)

Remark

The claim follows by the mathematical induction of $\max(p,q)$. *The cases* $p=1, q=2$, $p=2, q=1$ *and* $p=q=2$ *may be checked directly. Then, by separating a single block it is possible to prove the inductive step, i.e.,*

$$AB = \left[\begin{array}{c|ccc} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & & A_{2q} \\ \hline \vdots & & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right] \left[\begin{array}{c|ccc} B_{11} & B_{12} & \cdots & B_{1r} \\ \hline B_{21} & B_{22} & & B_{2r} \\ \hline \vdots & & \ddots & \vdots \\ \hline B_{q1} & B_{q2} & \cdots & B_{qr} \end{array}\right] =$$

$$= \left[\begin{array}{c|c} A_{1,1} & A'_{1,2} \\ \hline A'_{2,1} & A'_{2,2} \end{array}\right] \left[\begin{array}{c|c} B_{1,1} & B'_{1,2} \\ \hline B'_{2,1} & B'_{2,2} \end{array}\right] =$$

$$= \left[\begin{array}{c|c} A_{1,1}B_{1,1} + A'_{1,2}B'_{2,1} & A_{1,1}B'_{1,2} + A'_{1,2}B'_{2,2} \\ \hline A'_{1,2}B'_{1,1} + A'_{2,2}B'_{2,1} & A'_{2,1}B'_{2,2} + A'_{2,2}B'_{2,2} \end{array}\right] = C.$$

Multiplication of Block Matrices (continued)

Let $E_{ij} = [e_{kl}] \in M(m \times n; \mathbb{R})$ be a matrix such that

$$e_{kl} = \begin{cases} 1 & i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}.$$

Let δ_{ij} be the **Kronecker delta**, i.e.,

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Then, equivalently

$$e_{kl} = \delta_{ik}\delta_{jl},$$

and (matrices E, E' and E'' may have different sizes)

$$E_{ij}E'_{kl} = \delta jk E''_{il}.$$

In particular, for any matrix $A = [a_i j] \in M(m \times n; \mathbb{R})$

$$A = \sum_{i,j=1}^{n,m} a_{ij}E_{ij},$$

and the matrix multiplication can be seen as a special case of the block matrix multiplication.

Example

$$A = \left[\begin{array}{c|ccc} 1 & 2 & 2 & 6 \\ 1 & 2 & 2 & 5 \\ 1 & 1 & 2 & 8 \\ \hline 2 & 5 & 6 & 2 \end{array}\right], \quad B = \left[\begin{array}{c|cc} 1 & 0 & 2 \\ \hline 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right].$$

$$AB = \left[\begin{array}{c|c} \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1\end{bmatrix} + \begin{bmatrix}2&2&6\\2&2&5\\1&2&8\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix} & \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}0&2\end{bmatrix} + \begin{bmatrix}2&2&6\\2&2&5\\1&2&8\end{bmatrix}\begin{bmatrix}0&1\\1&2\\1&1\end{bmatrix} \\ \hline \begin{bmatrix}2\end{bmatrix}\begin{bmatrix}1\end{bmatrix} + \begin{bmatrix}5&6&2\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix} & \begin{bmatrix}2\end{bmatrix}\begin{bmatrix}0&2\end{bmatrix} + \begin{bmatrix}5&6&2\end{bmatrix}\begin{bmatrix}0&1\\1&2\\1&1\end{bmatrix} \end{array}\right]$$

$$= \left[\begin{array}{c|c} \begin{bmatrix}1\\1\\1\end{bmatrix} + \begin{bmatrix}6\\6\\5\end{bmatrix} & \begin{bmatrix}0&2\\0&2\\0&2\end{bmatrix} + \begin{bmatrix}8&12\\7&11\\10&13\end{bmatrix} \\ \hline 2+17 & \begin{bmatrix}0&4\end{bmatrix} + \begin{bmatrix}8&19\end{bmatrix} \end{array}\right] = \left[\begin{array}{c|cc} 7 & 8 & 14 \\ 7 & 7 & 13 \\ 6 & 10 & 15 \\ \hline 19 & 8 & 23 \end{array}\right].$$

Markov chains – Application

Let $S = \{1, \dots, N\}$ be the **state space**.

Definition

A (discrete–time, discrete–state, time–homogenous) **Markov chain**) is a sequence of random variables $X_0, X_1, X_2, \dots, X_n, \dots$ with values in the set S such that for all $i, i_0, \dots, i_{n-1}, j \in S$ and all $n \in \mathbb{N}$

$$P(X_{n+1} = j \mid X_n = i, X_{n-i} = i_{n-1}, \dots X_1 = i_1, X_0 = i_0) =$$

$$= P(X_{n+1} = j \mid X_n = i),$$

and it does not depend on n (i.e. the current state depends only on the previous state and this dependence is constant in time).

Markov chains – Application

Let $S = \{1, \ldots, N\}$ be the **state space**.

Definition

A (discrete–time, discrete–state, time–homogenous) **Markov chain**) is a sequence of random variables $X_0, X_1, X_2, \ldots, X_n, \ldots$ with values in the set S such that for all $i, i_0, \ldots, i_{n-1}, j \in S$ and all $n \in \mathbb{N}$

$$P(X_{n+1} = j \mid X_n = i, X_{n-i} = i_{n-1}, \ldots X_1 = i_1, X_0 = i_0) =$$

$$= P(X_{n+1} = j \mid X_n = i),$$

and it does not depend on n (i.e. the current state depends only on the previous state and this dependence is constant in time).

Remark

The number $P(X_{n+1} = j \mid X_n = i)$ is called the transition probability form the state i to the state j.

Markov chains – Application (continued)

Definition

The matrix $Q = [q_{ij}] \in M(N \times N; \mathbb{R})$ where

$$q_{ij} = P(X_{n+1} = j \mid X_n = i),$$

is called the **transition matrix**.

Markov chains – Application (continued)

Definition

The matrix $Q = [q_{ij}] \in M(N \times N; \mathbb{R})$ where

$$q_{ij} = P(X_{n+1} = j \mid X_n = i),$$

is called the **transition matrix**.

Example

For $N = 2$

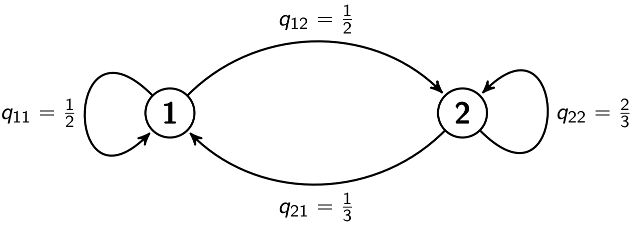
$$Q = \begin{bmatrix} P(X_{n+1} = 1 \mid X_n = 1) & P(X_{n+1} = 2 \mid X_n = 1) \\ P(X_{n+1} = 1 \mid X_n = 2) & P(X_{n+1} = 2 \mid X_n = 2) \end{bmatrix}.$$

Markov chains – Example

Each year a consumer of product A switches to product B with probability one–half. On the other hand, with probability two–thirds a consumer of product B continues buying it and with probability one–third starts buying product A. If $S = \{1, 2\}$ and 1 stands for product A and 2 for product B then

$$Q = \begin{bmatrix} P(X_{n+1} = 1 \mid X_n = 1) & P(X_{n+1} = 2 \mid X_n = 1) \\ P(X_{n+1} = 1 \mid X_n = 2) & P(X_{n+1} = 2 \mid X_n = 2) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Markov chains – Example (continued)



1 stands for product A and 2 stands for product B

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Markov chains – Example (continued)

What is the probability that a consumer of product A switches to product B after two years?

Markov chains – Example (continued)

What is the probability that a consumer of product A switches to product B after two years?

$$P(X_2 = 2 \mid X_0 = 1) =$$

$$= P(X_2 = 2 \mid X_1 = 1) P(X_1 = 1 \mid X_0 = 1) +$$
$$+ P(X_2 = 2 \mid X_1 = 2) P(X_1 = 2 \mid X_0 = 1) =$$
$$= \frac{1}{2} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4} + \frac{2}{6} = \frac{7}{12}.$$

(so a consumer switches either in the second or in the first year).

Markov chains (continued)

Definition

The n-step (conditional) probability of a Markov chain is

$$P(X_{n+2} = j \mid X_n = i) = q_{ij}^{(n)},$$

and the n-step condition matrix is

$$Q^{(n)} = [q_{ij}^{(n)}] \in M(N \times N; \mathbb{R}).$$

Markov chains (continued)

Definition

The n-step (conditional) probability of a Markov chain is

$$P(X_{n+2} = j \mid X_n = i) = q_{ij}^{(n)},$$

and the n-step condition matrix is

$$Q^{(n)} = [q_{ij}^{(n)}] \in M(N \times N; \mathbb{R}).$$

Proposition

$$Q^{(n)} = Q^n = \underbrace{Q \cdot Q \cdots Q}_{n\text{-times}}.$$

Markov chains (continued)

Proof.

It is enough to check the case $n = 2$. By the law of total probability

$$q_{ij}^{(2)} = P(X_{n+2} = j \mid X_n = i) =$$

$$= \sum_{s=1}^{N} P(X_{n+2} = j \mid X_{n+1} = s) P(X_{n+1} = s \mid X_n = i) =$$

$$= \sum_{s=1}^{N} q_{sj} q_{is} = \sum_{s=1}^{N} q_{is} q_{sj}.$$

$\square$

Example (continued)

If $Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$, then

$$Q^{(2)} = \begin{bmatrix} P(X_{n+2} = 1 \mid X_n = 1) & P(X_{n+2} = 2 \mid X_n = 1) \\ P(X_{n+2} = 1 \mid X_n = 2) & P(X_{n+2} = 2 \mid X_n = 2) \end{bmatrix} = Q^2 =$$

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \\ \frac{7}{18} & \frac{11}{18} \end{bmatrix}.$$

Initial Conditions and Marginal Distribution

Definition

The **initial conditions** is the (discrete) probability mass function of the variable X_0, i.e the vector

$$\mathbf{t} = (t_1, t_2, \ldots, t_N) =$$

$$= (P(X_0 = 1), P(X_0 = 2), \ldots, P(X_0 = N)) \in \mathbb{R}^N,$$

and the **marginal distributions** are the probability mass functions of variables $X_1, X_2, \ldots$

$$\mathbf{t}_i = (P(X_i = 1), P(X_i = 2), \ldots, P(X_i = N)) \in \mathbb{R}^N,$$

for $i = 1, 2, \ldots$.

Initial Conditions and Marginal Distribution

Definition

The **initial conditions** is the (discrete) probability mass function of the variable X_0, i.e the vector

$$\mathbf{t} = (t_1, t_2, \ldots, t_N) =$$

$$= (P(X_0 = 1), P(X_0 = 2), \ldots, P(X_0 = N)) \in \mathbb{R}^N,$$

and the **marginal distributions** are the probability mass functions of variables $X_1, X_2, \ldots$

$$\mathbf{t}_i = (P(X_i = 1), P(X_i = 2), \ldots, P(X_i = N)) \in \mathbb{R}^N,$$

for $i = 1, 2, \ldots$.

Remark

The marginal distributions depend on the initial conditions.

Initial Conditions and Marginal Distribution (continued)

Proposition

If Q is the transition matrix of a Markov chain then for $k \geqslant 1$

$$\mathbf{t}_k^\intercal = \mathbf{t}^\intercal Q^k.$$

Initial Conditions and Marginal Distribution (continued)

Proposition

If Q is the transition matrix of a Markov chain then for $k \geqslant 1$

$$\mathbf{t}_k^\intercal = \mathbf{t}^\intercal Q^k.$$

Proof.

By the law of total probability

$$P(X_k = i) = \sum_{s=1}^{N} P(X_k = i \mid X_0 = s) P(X_0 = s) =$$

$$= \sum_{s=1}^{N} t_s q_{si}^{(k)} =$$

$$= \text{the } i\text{-th entry of } \mathbf{t}^\intercal Q^k.$$

$\square$

Example

A consumer buys product A with probability $\frac{1}{5}$, product B with probability $\frac{4}{5}$ and the transition matrix is equal to $Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$,

so

$$\mathbf{t} = \left(\frac{1}{5}, \frac{4}{5}\right),$$

and

$$\mathbf{t}_2 = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \end{bmatrix} Q^2 = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \\ \frac{7}{18} & \frac{11}{18} \end{bmatrix} =$$

$$= \left(\frac{71}{180}, \frac{109}{180}\right).$$

Example (continued)

Hence, after 2 years, a consumer buys product A with probability $\frac{71}{180}$ and product B with probability $\frac{109}{180}$.

Stable State/Distribution

So, what happens after infinitely many years?

Stable State/Distribution

So, what happens after infinitely many years? Will the probability of buying product B be equal to one?

Stable State/Distribution

So, what happens after infinitely many years? Will the probability of buying product B be equal to one?

It turns out, in a distant time a consumer buys product A with probability $\frac{2}{5}$ and product B with probability $\frac{3}{5}$ (and the result does not depend on the initial conditions).

Stable State/Distribution

So, what happens after infinitely many years? Will the probability of buying product B be equal to one?

It turns out, in a distant time a consumer buys product A with probability $\frac{2}{5}$ and product B with probability $\frac{3}{5}$ (and the result does not depend on the initial conditions).

In particular the vector $\left(\frac{2}{5}, \frac{3}{5}\right)$ is a left eigenvector of matrix Q, or equivalently, an eigenvector of $Q^{\intercal}$, i.e.,

$$\begin{bmatrix} \frac{2}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{3} & \dfrac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

Hamming Codes

$(7, 4)$ Hamming code is a 2 error-detecting, 1 error-correcting linear code.

Hamming Codes

$(7,4)$ Hamming code is a 2 error-detecting, 1 error-correcting linear code.

$0000 \longrightarrow 0000000$	$1000 \longrightarrow 1110000$
$0001 \longrightarrow 1101001$	$1001 \longrightarrow 0011001$
$0010 \longrightarrow 0101010$	$1010 \longrightarrow 1011010$
$0011 \longrightarrow 1000011$	$1011 \longrightarrow 0110011$
$0100 \longrightarrow 1001100$	$1100 \longrightarrow 0111100$
$0101 \longrightarrow 0100101$	$1101 \longrightarrow 1010101$
$0110 \longrightarrow 1100110$	$1110 \longrightarrow 0010110$
$0111 \longrightarrow 0001111$	$1111 \longrightarrow 1111111$

$$b_0\, b_1\, b_2\, b_3 \rightarrow p_0\, b_0\, p_1\, p_2\, b_1\, b_2\, b_3.$$

Hamming Codes (continued)

$$b_0 b_1 b_2 b_3 \to p_0 b_0 p_1 p_2 b_1 b_2 b_3,$$

where

$$p_0 = b_0 + b_1 + b_3,$$

$$p_1 = b_0 + b_2 + b_3,$$

$$p_2 = b_0 + b_1 + b_3,$$

where addition is modulo 2, i.e $1 + 1 = 0$.

Hamming Codes (continued)

Encoding (and decoding) can be realised by matrix multiplication.

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} p_0 \\ b_0 \\ p_1 \\ p_2 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Paths in Directed Graphs

Definition

A **(simple, finite) directed graph** G is a pair $G = (V, E)$ where

$$V = \{v_1, \ldots, v_n\},$$

is the set of vertices and $E \subset V \times V$ is the set of edges (self–loops are allowed).

Paths in Directed Graphs

Definition

A **(simple, finite) directed graph** G is a pair $G = (V, E)$ where

$$V = \{v_1, \ldots, v_n\},$$

is the set of vertices and $E \subset V \times V$ is the set of edges (self–loops are allowed).

Definition

A **path** from $v \in V$ to $w \in V$ of **length** l is a sequence of edges

$$(v_{i_0}, v_{i_1}), (v_{i_1}, v_{i_2}), \ldots, (v_{i_{l-1}}, v_{i_l}),$$

such that

$$v_{i_0} = v, \quad v_{i_l} = w,$$

$$(v_{i_k}, v_{i_{k+1}}) \in E,$$

for $k = 0, \ldots, l - 1$.

Adjacency Matrix

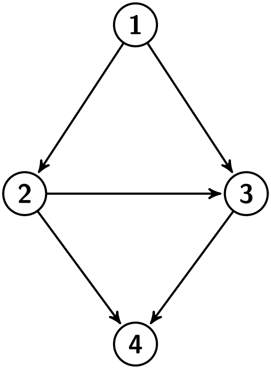
Definition

For a fixed simple, finite, directed graph G the **adjacency matrix** $A = A_G$ of G is the matrix $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ such that

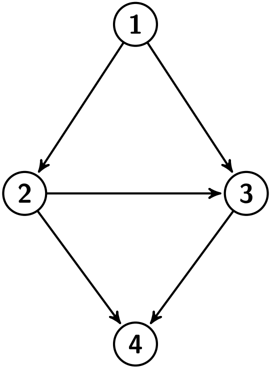
$$a_{ij} = 1 \text{ if and only if } (v_i, v_j) \in E,$$

and $a_{ij} = 0$ otherwise.

Example

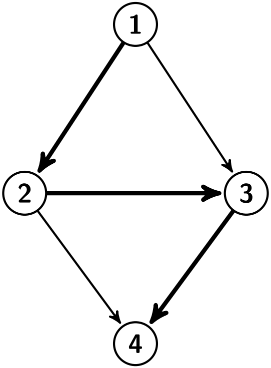


Example



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Example



For example $(v_1, v_2), (v_2, v_3), (v_3, v_4)$ is a path from v_1 to v_4 of length 3.

Number of Paths

Proposition

If A is the adjacency matrix of a (finite, simple) directed graph G, then $(A^l)_{ij}$ for any $l \geqslant 1$ is the number of paths of length l from v_i to v_j.

Number of Paths

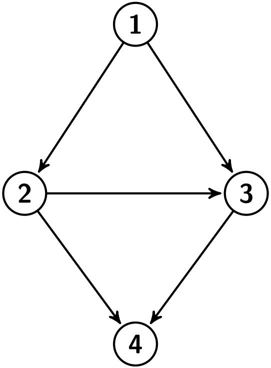
Proposition

If A is the adjacency matrix of a (finite, simple) directed graph G, then $(A^l)_{ij}$ for any $l \geqslant 1$ is the number of paths of length l from v_i to v_j.

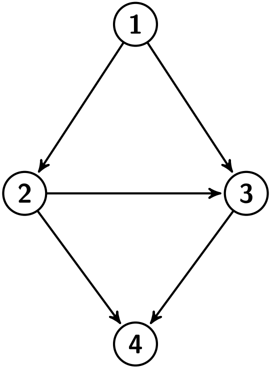
Proof.

By induction. $\square$

Example (continued)

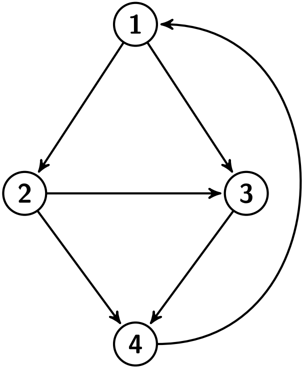


Example (continued)

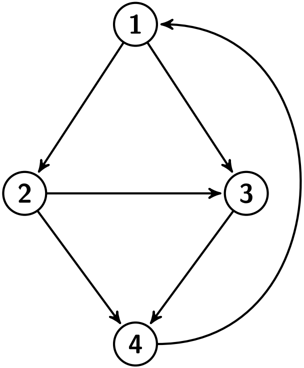


$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example



Example



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Dual Spaces

Definition

Let V be a vector space. The space dual to V is a vector space

$$V^* = \{\varphi\colon V \to \mathbb{R} \mid \varphi \text{ is linear}\} = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R}),$$

with the vector space structure given by

i) $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$,

ii) $(\alpha\varphi)v = \alpha\varphi(v)$,

for $\varphi, \psi \in V^*$ and any $v \in V$, $\alpha \in V$.

Dual Basis

Proposition

For any vector space V and any basis $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ of vector space V, there exists **the dual basis** $\mathcal{A}^* = (\alpha_1^*, \dots, \alpha_n^*)$ *of the vector space V^* such that*

$$\alpha_i^*(\alpha_j) = \begin{cases} 0 & \textit{for } i \neq j \\ 1 & \textit{for } i = j \end{cases}.$$

Proof.

There exists dual basis to the standard basis, i.e. st^* given by

$$\varepsilon_1^*((x_1, \dots, x_n)) = x_1,$$

$$\varepsilon_2^*((x_1, \dots, x_n)) = x_2,$$

$$\vdots$$

$$\varepsilon_n^*((x_1, \dots, x_n)) = x_n.$$

Dual Basis (continued)

Proof.

Let $\varphi\colon V \to V$ be the linear transformation such that

$$\varphi(\alpha_i) = \varepsilon_i.$$

Then

$$\alpha_i^* = \varepsilon_i^* \circ \alpha.$$

Assume $\alpha_i = (a_{1j}, a_{2j}, \ldots, a_{nj})$ and let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$, then

$$M(\varphi)_{st}^{st} = A^{-1}, \quad M(\varepsilon_i^*) = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$$

Dual Basis (continued)

Proof.

Let $A^{-1} = [b_{ij}]$. Then

$$M(\varepsilon_i)_{st}^{st} = M(\alpha_i^*)_{st}^{st} M(\varphi)_{st}^{st} =$$

$$= \begin{bmatrix} b_{i1} & b_{i2} & b_{i3} & \cdots & b_{in} \end{bmatrix}$$

That is

$$\varphi_i^* = b_{i1}\varepsilon_1^* + b_{i2}\varepsilon_2^* + \ldots + b_{in}\varepsilon_n^*.$$

$\square$

Example

Let $\alpha_1 = (1,2), \alpha_2 = (1,3)$ be a basis of $\mathbb{R}^2$. Then

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad B = A^{-1} = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}.$$

Therefore

$$\alpha_1^* = 3\varepsilon_1^* - \varepsilon_2^*,$$

$$\alpha_2^* = -2\varepsilon_1^* + \varepsilon_2^*,$$

or, in more concrete terms,

$$\alpha_1^*((x_1, x_2)) = 3x_1 - x_2,$$

$$\alpha_2^*((x_1, x_2)) = -2x_1 + x_2.$$

Dual of a linear transformation

Definition

For any linear transformation $\varphi \colon V \to W$, there exist the dual linear transformation

$$\varphi^* \colon W^* \to V^*,$$

given by the formula

$$\varphi^*(f)(v) = (f \circ \varphi)(v).$$

Dual of a linear transformation

Definition

For any linear transformation $\varphi\colon V \to W$, there exist the dual linear transformation

$$\varphi^*\colon W^* \to V^*,$$

given by the formula

$$\varphi^*(f)(v) = (f \circ \varphi)(v).$$

Proposition

If $\mathcal{A} = (\alpha_1, \ldots, \alpha_n), \mathcal{B} = (\beta_1, \ldots, \beta_m)$ *then*

$$M(\varphi^*)_{\mathcal{B}^*}^{\mathcal{A}^*} = \left[M(\varphi)_{\mathcal{A}}^{\mathcal{B}}\right]^{\mathsf{T}}.$$

Dual of a linear transformation (continued)

Proof.

Let $M(\varphi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}]$.

$$\varphi^*(\beta_i^*)(\alpha_j) = (\beta_i \circ \varphi)(\alpha_j) =$$

$$= \beta_i \left(\sum_{s=1}^{m} a_{sj} \beta_j \right) = a_{ij},$$

that the entry in the $i-$th column and in the $j-$th row of $M(\varphi^*)_{\mathcal{B}*}^{\mathcal{A}*}$ is equal to a_{ij}. $\square$

Dual subspaces

Definition

Let $W \subset V$ be any subset of vector space V. Let

$$W^* = \{f \in V^* \mid f\,|_W = 0\}.$$

If $W \subset V$ is a subspace the W^* called **the dual** subspace of W.

Dual subspaces

Definition

Let $W \subset V$ be any subset of vector space V. Let

$$W^* = \{f \in V^* \mid f\,|_W = 0\}.$$

If $W \subset V$ is a subspace the W^* called **the dual** subspace of W.

Proposition

Let $\varphi \colon V \to W$ be a linear transformation. Then

i) $\ker \varphi^* = (\operatorname{im} \varphi)^*$,

ii) $\operatorname{im} \varphi^* = (\ker \varphi)^*$.

Dual subspaces

Definition

Let $W \subset V$ be any subset of vector space V. Let

$$W^* = \{f \in V^* \mid f\mid_W = 0\}.$$

If $W \subset V$ is a subspace the W^* called **the dual** subspace of W.

Proposition

Let $\varphi\colon V \to W$ be a linear transformation. Then

i) $\ker \varphi^* = (\operatorname{im} \varphi)^*$,

ii) $\operatorname{im} \varphi^* = (\ker \varphi)^*$.

Proof.

Omitted. Needs isomorphism Theorem. $\square$

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B\colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B\colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

ii) $B(v, w + w') = B(v, w) + B(v, w')$ for any $v \in V, w, w' \in W$,

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B\colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

ii) $B(v, w + w') = B(v, w) + B(v, w')$ for any $v \in V, w, w' \in W$,

iii) $B(\alpha v, w) = \alpha B(v, w)$ for any $v \in V, w \in W, \alpha \in \mathbb{R}$,

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B\colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

ii) $B(v, w + w') = B(v, w) + B(v, w')$ for any $v \in V, w, w' \in W$,

iii) $B(\alpha v, w) = \alpha B(v, w)$ for any $v \in V, w \in W, \alpha \in \mathbb{R}$,

iv) $B(v, \beta w) = \beta B(v, w)$ for any $v \in V, w \in W, \beta \in \mathbb{R}$.

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B \colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

ii) $B(v, w + w') = B(v, w) + B(v, w')$ for any $v \in V, w, w' \in W$,

iii) $B(\alpha v, w) = \alpha B(v, w)$ for any $v \in V, w \in W, \alpha \in \mathbb{R}$,

iv) $B(v, \beta w) = \beta B(v, w)$ for any $v \in V, w \in W, \beta \in \mathbb{R}$.

Bilinear Forms and Bilinear Transformations

Definition

Let V, W, U be vector spaces. A function

$$B \colon V \times W \to U$$

is called a **bilinear transformation** if

i) $B(v + v', w) = B(v, w) + B(v', w)$ for any $v, v' \in V, w \in W$,

ii) $B(v, w + w') = B(v, w) + B(v, w')$ for any $v \in V, w, w' \in W$,

iii) $B(\alpha v, w) = \alpha B(v, w)$ for any $v \in V, w \in W, \alpha \in \mathbb{R}$,

iv) $B(v, \beta w) = \beta B(v, w)$ for any $v \in V, w \in W, \beta \in \mathbb{R}$.

Bilinear transformation B is called **a form** if moreover $U = \mathbb{R}$.

Equivalence Relation

Definition

An equivalence relation R on the set X is a relation (i.e. a subset) $R \subset X \times X$ such that

i) $\forall_{x \in X}\ (x, x) \in R$ (R is reflexive),

Equivalence Relation

Definition

An equivalence relation R on the set X is a relation (i.e. a subset) $R \subset X \times X$ such that

i) $\forall_{x \in X}\ (x, x) \in R$ (R is reflexive),

ii) $\forall_{x \in X} \forall_{y \in X}\ (x, y) \in R \Rightarrow (y, x) \in R$ (R is symmetric),

Equivalence Relation

Definition

An equivalence relation R on the set X is a relation (i.e. a subset) $R \subset X \times X$ such that

i) $\forall_{x \in X}\ (x, x) \in R$ (R is reflexive),

ii) $\forall_{x \in X} \forall_{y \in X}\ (x, y) \in R \Rightarrow (y, x) \in R$ (R is symmetric),

iii) $\forall_{x \in X} \forall_{y \in X} \forall_{z \in X}\ (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$ (R is transitive).

Partitions and Equivalence Classes

Equivalence relation R on set X induces a partition of X, given by its equivalence classes (in fact, there is a bijection between all partitions of X and all equivalence relations on X).

Definition

For any $x \in X$ the set

$$[x]_R = \{y \in X \mid (x, y) \in R\}$$

is called **an equivalence class** of element $x \in X$. When no confusion is possible we write $[x]$.

Quotient Vector Space

Definition

Let $W \subset V$ be a subspace of vector space V. The relation $R \subset V \times V$, given by the condition

$$(v, v') \in R \;\Leftrightarrow\; v - v' \in W,$$

is an equivalence relation on V, compatible with the structure of the vector space, i.e. the equivalence classes satisfy conditions

i) $[v] + [v'] = [v + v']$,

ii) $\alpha[v] = [\alpha v]$,

for any $v, v' \in V$ and $\alpha \in \mathbb{R}$.

Tensor Product

Definition

For any vector spaces V, W let U be a (infinite dimensional) vector space with basis $(u_{v,w})_{v\in V, w\in W}$ and let $U_0 \subset U$ be its subspace spanned by vectors

$$u_{v+v',w}-u_{v,w}-u_{v',w},\, u_{v,w+w'}-u_{v,w}-u_{v,w'},\, u_{\alpha v,w}-\alpha u_{v,w},\, u_{v,\beta w}-\beta u_{v,w},$$

where $v \in V, w \in V, \alpha, \beta \in \mathbb{R}$. By definition, **the tensor product** of vector spaces V and W is equal to the quotient space

$$V \otimes W = U/U_0.$$

Tensor Product (continued)

Definition

By definition

$$v \otimes w = [u_{v,w}].$$

Then, for any $v, v' \in V, w, w' \in W, \alpha, \beta \in \mathbb{R}$

i) $(v + v') \otimes w = v \otimes w + v' \otimes w,$

ii) $v \otimes (w + w') = v \otimes w + v \otimes w',$

iii) $(\alpha v) \otimes w = \alpha(v \otimes w),$

iv) $v \otimes (\beta w) = \beta(v \otimes w).$

Moreover, there exists bilinear transformation

$$\pi\colon V \times W \ni (v, w) \mapsto v \otimes w \in V \otimes W.$$

Tensor Product (continued)

Proposition

For any vector spaces V, W and for any bilinear transformation

$$B\colon V \times W \to U,$$

there exists **a unique linear transformation**

$$\varphi_B\colon V \otimes W \to U,$$

such that for any $v \in V, w \in W$

$$B(v, w) = \varphi_B(v \otimes w),$$

i.e.

$$\begin{tikzcd}
 & V \times W \arrow[dl, "\pi"'] \arrow[dr, "B"] & \\
V \otimes V \arrow[rr, dotted, "\varphi_B"] & & U
\end{tikzcd}$$

Tensor Product (continued)

Proof.

Let $\overline{\varphi}\colon \operatorname{lin}(\{u_{v,w} \mid v \in V,\ w \in W\}) \to U$ be a linear transformation given by

$$\overline{\varphi}(u_{v,w}) = B(v,w).$$

Since B is bilinear, the transformation $\overline{\varphi}$ sends vectors

$$u_{v+v',w} - u_{v,w} - u_{v',w},\ u_{v,w+w'} - u_{v,w} - u_{v,w'},\ u_{\alpha v,w} - \alpha u_{v,w},\ u_{v,\beta w} - \beta u_{v,w},$$

to zero (in U). Since vectors $[u_{v,w}]$ span $V \otimes W$, $\overline{\varphi}$ descends to a linear transformation

$$\varphi\colon V \otimes W \ni v \otimes w = [u_{v,w}] \mapsto B(v,w) \in U.$$

□

Tensor Product (continued)

Proposition

If $\mathcal{A} = (\alpha_1, \ldots \alpha_n)$ *is a basis of* V *and* $\mathcal{B} = (\beta_1, \ldots \beta_m)$ *is a basis of* W*, then*

$$\mathcal{A} \otimes \mathcal{B} = (\alpha_i \otimes \beta_j)_{i=1,\ldots,n, j=1,\ldots,m},$$

is a basis of $V \otimes W$*. It follows*

$$\dim V \otimes W = (\dim V)(\dim W) = nm.$$

Tensor Product (continued)

Proof.

Consider a bilinear form $B\colon V \times W \to \mathbb{R}$, i.e.

$$\begin{tikzcd} & V \times W \arrow[dl, "\pi"'] \arrow[dr, "B"] & \\ V \otimes W \arrow[rr, dotted, "\varphi_B"] & & \mathbb{R} \end{tikzcd}$$

If $v = \sum_{i=1}^{n} v_i \alpha_i$ and $w = \sum_{j=1}^{m} w_j \beta_j$ are bases of V and W, respectively, then

$$B(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{m} v_i w_j B(\alpha_i, \beta_j),$$

i.e. B is uniquely determined by the values $B(\alpha_i, \beta_j)$. On the other hand if $a_{ij} \in \mathbb{R}$ are some numbers, there exists a unique bilinear form B such that $B(\alpha_i, \beta_j) = a_{ij}$, given by the formula

$$B(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{m} v_i w_j a_{ij}.$$

Tensor Product (continued)

Proof.

By the properties of tensor product, any linear form φ induces a bilinear form $\pi \circ \varphi$.

$$\begin{tikzcd}
 & V \times W \arrow[dl, "\pi"'] \arrow[dr, "\varphi\circ\pi"] & \\
V \otimes W \arrow[rr, dotted, "\varphi"] & & \mathbb{R}
\end{tikzcd}$$

By the universal property, the linear transformation

$$(V \otimes W)^* \ni \varphi \mapsto \varphi \circ \pi,$$

is an isomorphism. Therefore

$$\dim V \otimes W = \dim(V \otimes W)^* = \dim V \dim W.$$

$\square$

Vector Space of Linear Transformations

Proposition

For any vector spaces V, W there exists a linear isomorphism of vector spaces

$$V^* \otimes W \ni f \otimes w \mapsto (V \ni v \mapsto f(v)w \in V) \in \mathrm{Hom}(V, W),$$

where $\mathrm{Hom}(V, W)$ *denotes the vector space of all linear transformations from V to W.*

Vector Space of Linear Transformations

Proposition

For any vector spaces V, W there exists a linear isomorphism of vector spaces

$$V^* \otimes W \ni f \otimes w \mapsto (V \ni v \mapsto f(v)w \in V) \in \mathrm{Hom}(V, W),$$

where $\mathrm{Hom}(V, W)$ denotes the vector space of all linear transformations from V to W.

Proof.

The transformation sends basis to basis. $\square$

Vector Space of Linear Transformations (continued)

Proposition

For any vector spaces V, W and for any $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ a basis of V and $\mathcal{B} = (\beta_1, \cdots, \beta_m)$ a basis of W, if

$$M(\varphi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}] \in M(m \times n; \mathbb{R}),$$

then, under the above isomorphism

$$\varphi = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \alpha_j^* \otimes \beta_i \in V^* \otimes W.$$

Vector Space of Linear Transformations (continued)

Proposition

For any vector spaces V, W and for any $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ a basis of V and $\mathcal{B} = (\beta_1, \cdots, \beta_m)$ a basis of W, if

$$M(\varphi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}] \in M(m \times n; \mathbb{R}),$$

then, under the above isomorphism

$$\varphi = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \alpha_j^* \otimes \beta_i \in V^* \otimes W.$$

Proof.

Exercise. □

Examples

If $\varphi\colon \mathbb{R}^3 \to \mathbb{R}$ is given by the formula

$$\varphi((x_1, x_2, x_3)) = x_1 + 2x_2 - 7x_3,$$

then

$$\varphi = \varepsilon_1^* + 2\varepsilon_2^* - 7\varepsilon_3^*.$$

Examples

If $\varphi\colon \mathbb{R}^3 \to \mathbb{R}$ is given by the formula

$$\varphi((x_1, x_2, x_3)) = x_1 + 2x_2 - 7x_3,$$

then

$$\varphi = \varepsilon_1^* + 2\varepsilon_2^* - 7\varepsilon_3^*.$$

If $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^2$, where $\varepsilon_1, \varepsilon_2, \varepsilon_3$ is the standard basis of $\mathbb{R}^3$ and ζ_1, ζ_2 is the standard basis of $\mathbb{R}^2$, is given by the formula

$$\varphi((x_1, x_2, x_3)) = (x_1 + 2x_2 - 7x_3, 5x_1 - 8x_2) =$$

$$x_1(1,5)+x_2(2,-8)+x_3(-7,0) = x_1(\zeta_1+5\zeta_2)+x_2(2\zeta_1-8\zeta_2)+x_3(-7\zeta_1),$$

then

$$M(\varphi)_{st}^{st} = \begin{bmatrix} 1 & 2 & -7 \\ 5 & -8 & 0 \end{bmatrix},$$

$$\varphi = \varepsilon_1^* \otimes \zeta_1 + 2\varepsilon_2^* \otimes \zeta_1 - 7\varepsilon_3^* \otimes \zeta_1 + 5\varepsilon_1^* \otimes \zeta_2 - 8\varepsilon_2^* \otimes \zeta_2.$$

Trace

Definition

Let $\varphi \colon V \to V$, i.e., $\varphi \in V^* \otimes V$ be a linear endomorphism. **The trace** of φ is equal to $\mathrm{Tr}(\varphi)$, where

$$\mathrm{Tr} \colon V^* \otimes V \ni \alpha \otimes \beta \mapsto \alpha(\beta) \in \mathbb{R},$$

is the unique linear transformation corresponding to the bilinear transformation

$$V^* \times V \ni (\alpha, \beta) \mapsto \alpha(\beta) \in \mathbb{R}.$$

Trace

Definition

Let $\varphi\colon V \to V$, i.e., $\varphi \in V^* \otimes V$ be a linear endomorphism. **The trace** of φ is equal to $\mathrm{Tr}(\varphi)$, where

$$\mathrm{Tr}\colon V^* \otimes V \ni \alpha \otimes \beta \mapsto \alpha(\beta) \in \mathbb{R},$$

is the unique linear transformation corresponding to the bilinear transformation

$$V^* \times V \ni (\alpha, \beta) \mapsto \alpha(\beta) \in \mathbb{R}.$$

Proposition

For any basis $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ *of* V*, if* $M(\varphi)^{\mathcal{A}}_{\mathcal{A}} = [a_{ij}]$*, then*

$$\mathrm{Tr}(\varphi) = \sum_{i=1}^{n} a_{ii}.$$

Trace (continued)

Proof.

By the definition

$$\varphi = \sum_{i,j=1}^{n} a_{ij}\alpha_j^* \otimes \alpha_i,$$

and

$$\mathrm{Tr}\left(\alpha_j^* \otimes \alpha_i\right) = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j, \end{cases}.$$

□

Trace of a Matrix

Definition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix. **The trace** of matrix A is equal to

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Trace of a Matrix

Definition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix. **The trace** of matrix A is equal to

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Corollary

For any linear endomorphism $\varphi \colon V \to V$ and any basis $\mathcal{A}$ of vector space V

$$\mathrm{Tr}(\varphi) = \mathrm{Tr}\big(M(\varphi)_{\mathcal{A}}^{\mathcal{A}}\big),$$

that is for any invertible matrix $C \in M(n \times n; \mathbb{R})$

$$\mathrm{Tr}(A) = \mathrm{Tr}\big(C^{-1}AC\big),$$

i.e., the trace admits the same value on similar matrices.

Trace of a Matrix (continued)

Proposition

For any matrices $A, B, C \in M(n \times n; \mathbb{R})$ *and scalars* $\alpha, \beta \in \mathbb{R}$

i) $\operatorname{Tr}(A) = \operatorname{Tr}(A^\intercal)$,

ii) $\operatorname{Tr}(\alpha A + \beta B) = \alpha \operatorname{Tr}(A) + \beta \operatorname{Tr}(B)$,

iii) $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$,

iv) $\operatorname{Tr}(ABC) = \operatorname{Tr}(BCA) = \operatorname{Tr}(CAB)$.

Proof.

Points i) and ii) are obvious and iv) follows from iii).

Trace of a Matrix (continued)

Proof.

iii) let $C = [c_{ij}] = AB$ and $C' = [c'_{ij}] = BA$. Then

$$\mathrm{Tr}(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ji},$$

$$\mathrm{Tr}\big(C'\big) = \sum_{i=1}^{n} c'_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} a_{ji} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ji} b_{ij} = \mathrm{Tr}(C).$$

□

Non–canonical Isomorphism

Remark

If V is a finite dimensional vector space then the vector spaces V and V^* are isomorphic, however there is no canonical isomorphism.

Example

For example, let $V = \mathbb{R}^2$ and let $F, G\colon V \to V^*$ be linear transformations given by the conditions

$$F(\varepsilon_1) = \varepsilon_1^*, \quad G(\alpha_1) = \alpha_1^*,$$

$$F(\varepsilon_2) = \varepsilon_2^*, \quad G(\alpha_2) = \alpha_2^*,$$

where $\alpha_1 = (1,2), \alpha_2 = (1,3)$ then

$$F(\alpha_1) = F(\varepsilon_1 + 2\varepsilon_2) = \varepsilon_1^* + 2\varepsilon_2^* \neq \alpha_1^* =$$

$$= 3\varepsilon_1^* - \varepsilon_2^*.$$

Non–canonical Isomorphism

Remark

If V is infinite dimensional then V is not isomorphic to V^* (the reason is purely set–theoretical). For example

$$(\mathbb{R} \oplus \mathbb{R} \oplus \dots)^* \simeq (\mathbb{R} \times \mathbb{R} \times \dots).$$

The Bidual Space

Definition

For any vector space V, the **bidual space** is equal to

$$V^{**} = (V^*)^*.$$

The Bidual Space

Definition

For any vector space V, the **bidual space** is equal to

$$V^{**} = (V^*)^*.$$

Proposition

If V is finite dimensional, there exists a canonical isomorphism of V and V^{**} (which does not depend on the particular choice of a basis)

$$F\colon V \ni v \mapsto (V^* \ni f \mapsto f(v) \in \mathbb{R}) \in (V^*)^* = V^{**}.$$

The Bidual Space

Definition

For any vector space V, the **bidual space** is equal to

$$V^{**} = (V^*)^*.$$

Proposition

If V is finite dimensional, there exists a canonical isomorphism of V and V^{**} (which does not depend on the particular choice of a basis)

$$F\colon V \ni v \mapsto (V^* \ni f \mapsto f(v) \in \mathbb{R}) \in (V^*)^* = V^{**}.$$

Proof.

$$v \in \ker F \Leftrightarrow f(v) = 0 \text{ for all } f \in V^* \Leftrightarrow v = \mathbf{0},$$

since any non–zero vector can be completed to basis, and then $v^*(v) = 1$. $\square$

Frobenius Norm

Proposition

The bilinear (real, Frobenius) form

$$\langle\cdot,\cdot\rangle\colon M(m\times n;\mathbb{R})\times M(m\times n;\mathbb{R})\ni(A,B)\mapsto\mathrm{Tr}(A^\intercal B)\in\mathbb{R},$$

is

i) *symmetric, i.e.,*
$$\langle A,B\rangle=\langle B,A\rangle,$$

ii) *positive definite, i.e.,*
$$\langle A,A\rangle>0,$$

if $A\neq\mathbf{0}$*,*

Frobenius Norm (continued)

Proposition

iii) *non–degenerate, i.e., the linear transformation*

$$\langle A, \cdot \rangle \colon M(m \times n; \mathbb{R}) \ni B \mapsto \langle A, B \rangle \in \mathbb{R},$$

is non–zero if and only if $A \neq \mathbf{0}$,

iv) *invariant under the left and the right multiplication by an orthogonal matrix, i.e., if* $Q \in M(m \times m; \mathbb{R})$ *satisfies* $Q^\intercal Q = QQ^\intercal = I_m$, *then*

$$\langle QA, QB \rangle = \langle A, B \rangle,$$

and if $P \in M(n \times n; \mathbb{R})$ *satisfies* $P^\intercal P = PP^\intercal = I_n$, *then*

$$\langle AP, BP \rangle = \langle A, B \rangle.$$

Frobenius Norm (continued)

Proof.

i) $\mathrm{Tr}(A^\intercal B) = \mathrm{Tr}((A^\intercal B)^\intercal) = \mathrm{Tr}(B^\intercal A)$,

ii) if $A = [a_{ij}] \neq \mathbf{0}$, then
$\mathrm{Tr}(A^\intercal A) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} a_{ji} = \sum_{i,j=1}^{n} a_{ij}^2 > 0$,

iii) follows from *ii*) (substitute $B = A$),

iv) $\mathrm{Tr}((QA)^\intercal QB) = \mathrm{Tr}(A^\intercal Q^\intercal QB) = \mathrm{Tr}(A^\intercal B)$ and
$\mathrm{Tr}((AP)^\intercal BP) = \mathrm{Tr}(P^\intercal A^\intercal BP) = \mathrm{Tr}(A^\intercal B)$ since $P^\intercal = P^{-1}$.

$\square$

Frobenius Norm (continued)

Definition

For any matrix $A \in M(m \times n; \mathbb{R})$ the **Frobenius norm** of A is equal to

$$\|A\|_F = \sqrt{\langle A, A\rangle} = \sqrt{\operatorname{Tr}(A^\intercal A)}.$$

Frobenius Norm (continued)

Proposition

i) $\|A\|_F = 0 \quad \Leftrightarrow \quad A = \mathbf{0}$,

ii) $\|\alpha A\|_F = |\alpha| \|A\|_F$ *for any scalar* $\alpha \in \mathbb{R}$,

iii) $\|A + B\|_F \leqslant \|A\|_F + \|B\|_F$, *for any matrices* $A, B \in M(n \times m; \mathbb{R})$

iv) $\|A\|_F = \|A^\intercal\|_F$,

v) $\|AB\|_F \leqslant \|A\|_F \|B\|_F$, *for any matrix* $A \in M(m \times n; \mathbb{R})$ *and any matrix* $B \in M(n \times k; \mathbb{R})$,

vi) $\|QAP\|_F = \|A\|_F$ *for any matrix* $A \in M(n \times m; \mathbb{R})$, *any orthogonal matrix* $Q \in M(m \times m; \mathbb{R})$ *and any orthogonal matrix* $P \in M(n \times n; \mathbb{R})$

Frobenius Norm (continued)

Proof.

i) the Frobenius form is positive definite,

ii) obvious,

iii) by the Cauchy–Schwarz inequality for the Frobenius form $|\langle A, B\rangle| \leqslant \|A\|_F \|B\|_F$, then $\|A+B\|_F^2 = \|A\|_F^2 + 2\langle A, B\rangle + \|B\|_F^2 \leqslant (\|A\|_F + \|B\|_F)^2$,

iv) obvious,

$\square$

Frobenius Norm (continued)

Proof.

v) from the Cauchy–Schwarz inequality for the standard scalar product in $\mathbb{R}^n$, if $AB = [c_{pq}]$,

$$\|AB\|_F^2 = \sum_{p=1}^{m} \sum_{q=1}^{k} c_{pq}^2 =$$

$$\sum_{p=1}^{m} \sum_{q=1}^{k} \left(\sum_{r=1}^{n} a_{pr} b_{rq} \right)^2 \leqslant \sum_{p=1}^{m} \sum_{q=1}^{n} \left(\sum_{r=1}^{n} a_{pr}^2 \right) \left(\sum_{r'=1}^{n} b_{qr'}^2 \right) =$$

$$= \left(\sum_{p=1}^{m} \left(\sum_{r=1}^{n} a_{pr}^2 \right) \right) \left(\sum_{q=1}^{k} \left(\sum_{r'=1}^{n} b_{qr'}^2 \right) \right) = \|A\|_F^2 \|B\|_F^2.$$

vi) follows from the properties of the Frobenius form.

$\square$

Frobenius Norm (continued)

Remark

$\|AB\|_F = \|A\|_F \|B\|_F$ holds if and only if $i-th$ row of A is linearly dependent with $i-th$ column of B for any i (equality in standard Cauchy–Schwarz inequality).

$\|A+B\|_F = \|A\|_F + \|B\|_F$ holds if and only if $\langle A, B\rangle = \|A\|\|B\|$, i.e. $A = \lambda B$ or $B = \lambda A$ for some $\lambda \geqslant 0$ (as in the standard Minkowski inequality).

Frobenius Norm (continued)

Remark

$\|AB\|_F = \|A\|_F \|B\|_F$ *holds if and only if* $i-$*th row of* A *is linearly dependent with* $i-$*th column of* B *for any* i *(equality in standard Cauchy–Schwarz inequality).*

$\|A+B\|_F = \|A\|_F + \|B\|_F$ *holds if and only if* $\langle A, B\rangle = \|A\|\|B\|$*, i.e.* $A = \lambda B$ *or* $B = \lambda A$ *for some* $\lambda \geqslant 0$ *(as in the standard Minkowski inequality).*

By the SVD decomposition if $\sigma_1(A), \ldots, \sigma_r(A) \in \mathbb{R}$ *denote the singular values of* A *(i.e., square roots of the non–zero eigenvalues of the matrix* $A^\intercal A$*) then*

$$\|A\|_F = \sqrt{\sigma_1^2(A) + \ldots + \sigma_r^2(A)}.$$

Frobenius Norm (continued)

$$\|A\|_F = \sqrt{\sigma_1^2(A) + \ldots + \sigma_r^2(A)},$$

$$\|A\|_2 = \sigma_1(A),$$

therefore for any matrix $A \in M(m \times n; \mathbb{C})$

$$\|A\|_2 \leqslant \|A\|_F \leqslant \sqrt{\min\{m, n\}}\|A\|_2.$$

Note that any two norms in a finite–dimensional space are equivalent (i.e. they induce the same topology and convergence in one norm is equivalent to the convergence in the other).

Matrix Multiplication as a Sum of Rank 1 Matrices

Remark

In some contexts (large sparse matrices) it is useful to interpret the matrix product in the following way. Let $A \in M(m \times n; \mathbb{R})$ and let $B \in M(n \times l; \mathbb{R})$. The matrix product of A by B is equal to the sum rank 1 matrices

$$AB = \sum_{s=1}^{n} C_i R_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \begin{bmatrix} b_{i1} & b_{i2} & \dots & b_{in} \end{bmatrix}.$$

Proof.

The (i, j) entry of $C_s R_s$ is equal to $a_{is} b_{sj}$. □